



General broken lines as advanced track fitting method

Claus Kleinwort*

DESY Deutsches Elektronen-Synchrotron, Notkestraße 85, 22607 Hamburg, Germany

ARTICLE INFO

Article history:

Received 13 July 2011

Received in revised form

13 January 2012

Accepted 13 January 2012

Available online 21 January 2012

Keywords:

Track fitting

Multiple scattering

Global least-squares fit

Millepede

Broken lines

ABSTRACT

In HEP experiments the description of the trajectory of a charged particle is obtained from a fit to measurements in tracking detectors. The parameterization of the trajectory has to account for bending in the magnetic field, energy loss and multiple scattering in the detector material. General broken lines implement a track model with the proper description of multiple scattering leading to linear with a special structure of the corresponding matrix allowing for a fast solution with the computing time depending linearly on the number of measurements. The calculation of the full covariance matrix along the trajectory enables the application to track based alignment and calibration of large detectors with global methods.

© 2012 Elsevier B.V. All rights reserved.

1. Introduction

The trajectory of a charged particle in a homogeneous magnetic field neglecting the interactions with the detector material is described by a helix. In a global coordinate system (x, y, z) with the magnetic field in z -direction it can be parameterized by the inverse momentum (times charge) q/p , an angle ϕ_0 at and the distance d_0 to the point of the closest approach in the (x, y) -plane, the dip angle λ to that plane and the offset z_0 at the point of the closest approach. Energy loss in the detector material due to ionization or radiation (for electrons) leads to a reduction of the momentum. *Multiple scattering*, mainly due to Coulomb interaction with the electrons in the atoms, results in random changes in direction and spatial position with expectation values of zero and variance depending on the traversed material and the particle momentum. In addition the magnetic field is usually homogeneous only in an approximation. Therefore more advanced track models are required.

The effect of multiple scattering can be taken into account in different ways [1]. For global methods it can be added to the measurement errors leading to a non-diagonal covariance matrix or explicitly fitted by scattering angles as additional track parameters. In both cases this requires the inversion of a large matrix of order n (number of measurements or number of scatterers) with computing time $\mathcal{O}(n^3)$. Progressive methods [2] like the Kálmán filter [3] updating the track parameters for each additional measurement and scatterer with a computing time $\mathcal{O}(n)$ have become the standard method for track fitting.

The broken lines' method is a fast global track refit adding the description of multiple scattering to an initial trajectory and able to determine the complete covariance matrix of all track parameters. This allows the usage as track model for track based alignment and calibration with the global Millepede-II [4] algorithm. Corrections and covariance matrices for the local track parameters (at single points) can be determined with a computing time $\mathcal{O}(n)$ exploiting the sparsity of the matrix of the corresponding linear equations system.

The original formulation [5] describes the case of a tracking detector with a solenoidal magnetic field and independent two-dimensional tracking in the bending and perpendicular to the bending plane. It constructs the planar trajectories from the measurements including the material around those as thick scatterers. In the presence of measurements with components in both planes a single trajectory in three dimensions is required. General broken lines are describing trajectories in space in an arbitrary magnetic field for arbitrary measurements and material distribution. The three-dimensional broken lines' trajectories from one- or two-dimensional independent measurements dedicated to the track-based alignment of the CMS silicon tracker [6] with Millepede-II [4] are an intermediate step. They are constructed from the material between the measurement planes by equivalent thin scatterers. Measurements are parameterized by the interpolation of (small) offsets perpendicular to the track direction defined at those scatterers and a common correction $\Delta q/p$ of the inverse momentum. Multiple scattering kinks are defined by triplets of thin scatterers. The corresponding linear equation system $\mathbf{Ax} = \mathbf{b}$ with a bordered band matrix \mathbf{A} with a size n depending on the number of scatterers n_{scat} , band width m (usually 5) and border size $b = 1$ is built from the Jacobians of the

* Tel.: +49 4089983545; fax: +49 40 8994 3545.

E-mail address: Claus.Kleinwort@desy.de

propagation of the offsets between measurements and scatterers on the initial trajectory and solved by root-free Cholesky decomposition. In the following the further features of general broken lines are presented, the procedure is discussed and a comparison with the Kálmán filter is performed.

2. Towards general broken lines

2.1. Local coordinate system

At each measurement plane and each thin scatterer a local (orthonormal) coordinate system (u_1, u_2, w) is defined. The natural choice for the w -direction is perpendicular to the measurement plane for a measurement and parallel to the track direction for a scatterer. A local system moving with the track is the *curvilinear* frame $(x_\perp, y_\perp, z_\perp)$ with z_\perp in track direction and x_\perp in the global (x, y) -plane. At each thin scatterer the two-dimensional offset $\mathbf{u} = (u_1, u_2)$ is defined as fit parameter.

The variance \mathbf{V}_k of the multiple scattering kinks \mathbf{k} in the local system can be calculated from the variance $\theta_0^2 \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ [7] in the curvilinear system by parameter transformation [8] from the curvilinear to the local slopes which depends on the projections $c_i = \mathbf{e}_{\text{track}} \cdot \mathbf{e}_{u_i}$ of the offset directions \mathbf{e}_{u_i} of the local frame onto the track direction $\mathbf{e}_{\text{track}}$:

$$\mathbf{V}_k = \frac{\partial(u'_1, u'_2)}{\partial(x'_\perp, y'_\perp)} \begin{pmatrix} \theta_0^2 & 0 \\ 0 & \theta_0^2 \end{pmatrix} \left[\frac{\partial(u'_1, u'_2)}{\partial(x'_\perp, y'_\perp)} \right]^T$$

$$= \frac{\theta_0^2}{(1 - c_1^2 - c_2^2)^2} \begin{pmatrix} 1 - c_2^2 & c_1 c_2 \\ c_1 c_2 & 1 - c_1^2 \end{pmatrix} \quad (1)$$

With at least one offset defined perpendicular to the track direction this is a diagonal matrix.

2.2. Local track parameters

In the local frame a track can be described by an offset (u_1, u_2) , a slope $(u'_1, u'_2) = \partial(u_1, u_2)/\partial w$ and the inverse momentum q/p . The general broken lines' fit determines from the fit parameters $\mathbf{x} = (\Delta q/p, \mathbf{u}_1, \dots, \mathbf{u}_{n_{\text{scat}}})$ at each thin scatterer i track parameter corrections $\Delta \mathbf{p}_{\text{loc}} = (\Delta q/p, \mathbf{u}'_i, \mathbf{u}_i)$. The slope before the scattering is defined by $(\Delta q/p, \mathbf{u}_{i-1}, \mathbf{u}_i)$ and afterwards by $(\Delta q/p, \mathbf{u}_i, \mathbf{u}_{i+1})$ (Eq. (6) in Ref. [6]). At the measurement planes the track parameter corrections $\Delta \mathbf{p}_{\text{loc}} = (\Delta q/p, \mathbf{u}'_{\text{int}}, \mathbf{u}_{\text{int}})$ can be obtained from the interpolation of the enclosing scatterers (Eq. (8) and inserting \mathbf{u}_{int} for \mathbf{u}_0 in Eq. (6) in Ref. [6]). In any case $\Delta \mathbf{p}_{\text{loc}}$ depends only on $\Delta q/p$ and two adjacent offsets. Therefore for the covariance matrices of the track parameter corrections the bordered band part of the covariance matrix \mathbf{A}^{-1} of the fit parameters is sufficient. For a sparse matrix the elements of the inverse matrix inside the sparsity pattern can be calculated by special methods without those outside that pattern [9]. This allows to obtain the bordered band part of \mathbf{A}^{-1} with $\mathcal{O}(n(m+b)^2)$ operations linear in the number of scatterers.

2.3. Measurements

The measurements \mathbf{m} are (the residuals with respect to the initial trajectory of) arbitrary observables with predictions defined by the local track parameters in the measurement plane. This can be one- or two-dimensional position measurements or track segments containing slope information. The linearized prediction is $\mathbf{m}(\mathbf{x}) = \mathbf{H}_m \mathbf{x}$. In case the corresponding covariance matrices \mathbf{V}_m are not diagonal they should be diagonalized and the measurements and predictions modified accordingly. This allows the usage of univariate M-estimators for outlier down-weighting and the interfacing to Millepede-II [4] expecting independent scalar measurements.

2.4. Iterations

In case the resulting fit parameters are not small corrections as assumed for the linearization of the track model (Eq. (3) in Ref. [6]) the trajectory has to be iterated. This requires initial non-zero values \mathbf{k}_0 in the prediction of the multiple scattering kinks: $\mathbf{k}(\mathbf{x}) = \mathbf{H}_k \mathbf{x} + \mathbf{k}_0$. The expectation value remains $\langle \mathbf{k}(\mathbf{x}) \rangle = \mathbf{0}$.

2.5. Seeding

The general broken lines are seeded by an initial trajectory. Alternatively to the result of a fit of the measurements (*internal seeding*) this can be a prediction from another part of the detector (*external seeding*). The seeding track parameters at some reference point are used for the propagation along the trajectory according to the magnetic field (and average energy loss) to calculate residuals for the measurements and parameter transformation matrices. If the track fit has to be iterated to account for non-linear effects the initial trajectory could be based on general broken lines itself. On average no change of the track parameters at the reference point is expected from refitting: $\langle \Delta \mathbf{p}_{\text{seed}} \rangle = \langle \mathbf{H}_s \mathbf{x} \rangle = \mathbf{0}$.

2.6. Minimization

The fit parameters are determined by minimizing

$$\chi^2(\mathbf{x}) = \sum_{i=1}^{n_{\text{meas}}} (\mathbf{H}_{m,i} \mathbf{x} - \mathbf{m}_i)^T \mathbf{V}_{m,i}^{-1} (\mathbf{H}_{m,i} \mathbf{x} - \mathbf{m}_i) \quad (\text{from measurements})$$

$$+ \sum_{i=2}^{n_{\text{scat}}-1} (\mathbf{H}_{k,i} \mathbf{x} + \mathbf{k}_{0,i})^T \mathbf{V}_{k,i}^{-1} (\mathbf{H}_{k,i} \mathbf{x} + \mathbf{k}_{0,i}) \quad (\text{from kinks})$$

$$+ (\mathbf{H}_s \mathbf{x})^T \mathbf{V}_s^{-1} (\mathbf{H}_s \mathbf{x}) \quad (\text{from external seed}) \quad (2)$$

The minimization leads to a linear equation system $\mathbf{A} \mathbf{x} = \mathbf{b}$ with \mathbf{x} of size $n = 2n_{\text{scat}} + 1$:

$$\mathbf{A} = \sum_{i=1}^{n_{\text{meas}}} \mathbf{H}_{m,i}^T \mathbf{V}_{m,i}^{-1} \mathbf{H}_{m,i} + \sum_{i=2}^{n_{\text{scat}}-1} \mathbf{H}_{k,i}^T \mathbf{V}_{k,i}^{-1} \mathbf{H}_{k,i} + \mathbf{H}_s^T \mathbf{V}_s^{-1} \mathbf{H}_s \quad (3a)$$

$$\mathbf{b} = \sum_{i=1}^{n_{\text{meas}}} \mathbf{H}_{m,i}^T \mathbf{V}_{m,i}^{-1} \mathbf{m}_i - \sum_{i=2}^{n_{\text{scat}}-1} \mathbf{H}_{k,i}^T \mathbf{V}_{k,i}^{-1} \mathbf{k}_{0,i} \quad (3b)$$

3. Procedure

The initial trajectory provides in the local coordinate system for the measurements the residuals \mathbf{m} with (diagonal) covariance matrix \mathbf{V}_m , for the scatterers the initial (usually zero) kinks \mathbf{k}_0 with covariance matrix \mathbf{V}_k , optionally the external seed and for the parameter transformations from one point (measurement, scatterer or seed) to the next one the Jacobians $\mathbf{T}_i = \partial \mathbf{p}_{\text{loc},i+1} / \partial \mathbf{p}_{\text{loc},i}$.

The linearized track model of the general broken lines is defined at each point by the offset part $(\mathbf{d}, \mathbf{S}, \mathbf{J})_\pm = \partial \mathbf{u} / \partial (q/p, \mathbf{u}, \mathbf{u})_\pm$ of the Jacobians for the transformation to the next and the previous scatterer. For the forward propagation they are calculated from the products of the point-to-point Jacobians \mathbf{T} , for the backward case additional partial inversions (for the offset part) are performed. In the next step the non-zero elements of the matrices \mathbf{H} for the linearized predictions are calculated from the 2×2 matrices and vectors \mathbf{J}_\pm , $\mathbf{W}_\pm = \pm \mathbf{S}_\pm^{-1}$ and \mathbf{d}_\pm (see Eqs. (7) and (8) in Ref. [6]). The measurements \mathbf{m} , the kinks \mathbf{k} , the derivatives from the sparse matrices \mathbf{H} and the measurement and multiple scattering errors from diagonal covariance matrices \mathbf{V} are the ingredients for the local fit in Millepede-II [4]. Afterwards the linear equation system $\mathbf{A} \mathbf{x} = \mathbf{b}$ is constructed from the individual contributions $(\mathbf{H}^T \mathbf{V}^{-1} \mathbf{H}, \mathbf{H}^T \mathbf{V}^{-1} \mathbf{m}$ or $\mathbf{H}^T \mathbf{V}^{-1} \mathbf{k}_0)$ of

measurements, kinks and optionally the external seed. The band part \mathbf{A}_u of the matrix is decomposed into a diagonal matrix \mathbf{D} and a triangular band matrix \mathbf{L} with unit diagonal ($\mathbf{A}_u = \mathbf{LDL}^T$) and the offsets \mathbf{x}_u are determined from \mathbf{D} , \mathbf{L} and \mathbf{b}_u by solving $\mathbf{Lz} = \mathbf{b}_u$ and $\mathbf{L}^T \mathbf{x}_u = \mathbf{D}^{-1} \mathbf{z}$ [5]. The band part of \mathbf{A}_u^{-1} is calculated according to Ref. [9]. Using block matrix algebra the complete solution \mathbf{x} and the bordered band part of \mathbf{A}^{-1} are obtained. From this corrections and covariance matrices for local track parameters at any point (measurement, scatterer or seed) can be derived.

For all this steps (propagation, prediction, construction, decomposition, solution and inversion) the number of operations is linear in the number of points or parameters and a sizable contribution to the total computing time. As all predictions \mathbf{Hx} depend at most on three consecutive offsets the matrix \mathbf{A}_u has a band width $m = 3 \cdot 2 - 1 = 5$. In case the transformation matrices \mathbf{J} and \mathbf{S} are diagonal the two offset components are decoupled and the band width reduces to $m = (3-1) \cdot 2 = 4$ speeding up the band matrix operations ($\mathcal{O}(n(m+b)^2)$). In addition the construction of \mathbf{A} is faster due to fewer non-zero elements in \mathbf{H} .

4. Comparison with Kálmán filter

4.1. Mathematical equivalence

An externally seeded general broken lines trajectory with one measurement has only one offset defined and the slope has to be used directly as fit parameter: $\mathbf{x} = \Delta \mathbf{p}_{\text{loc},1} = (\Delta q/p, \mathbf{u}_1, \mathbf{u}_1)$. The

solution of the linear equation system is

$$\mathbf{x} = \mathbf{A}^{-1} [\mathbf{H}_{m,1}^T \mathbf{V}_{m,1}^{-1} \mathbf{m}_1] \quad (4a)$$

$$\mathbf{A}^{-1} = [\mathbf{V}_s^{-1} + \mathbf{H}_{m,1}^T \mathbf{V}_{m,1}^{-1} \mathbf{H}_{m,1}]^{-1} \quad (4b)$$

This is equivalent to the filtering step of a Kálmán filter in the weighted mean formalism (Eq. (8b) in Ref. [3]):

$$\mathbf{x}_k = \mathbf{C}_k [(\mathbf{C}_k^{k-1})^{-1} \mathbf{x}_k^{k-1} + \mathbf{H}_k^T \mathbf{V}_k^{-1} \mathbf{m}_k] \quad (5a)$$

$$\mathbf{C}_k = [(\mathbf{C}_k^{k-1})^{-1} + \mathbf{H}_k^T \mathbf{V}_k^{-1} \mathbf{H}_k]^{-1} \quad (5b)$$

As the initial trajectory has been based on the track parameters from the external seed the prediction \mathbf{x}_k^{k-1} is zero, \mathbf{C}_k corresponds to \mathbf{A}^{-1} , \mathbf{C}_k^{k-1} to \mathbf{V}_s and \mathbf{m}_k to \mathbf{m}_1 .

In general only the covariance matrices for the track parameters at single points are calculated by the Kálmán filter. For the global covariance matrix from all track parameters of the trajectory required by global alignment and calibration methods an extension of the Kálmán filter is described in Ref. [10].

4.2. Computational difference

Track fitting with the Kálmán filter algorithm [3] is a sequential procedure adding measurements and scatterers (process noise) to the trajectory one at the time. The propagation of track parameters and covariance matrices from point to point in the filtering step involves multiplications of 5×5 matrices with $\mathcal{O}(5^3)$ operations. The filtering in the weighted mean formalism is

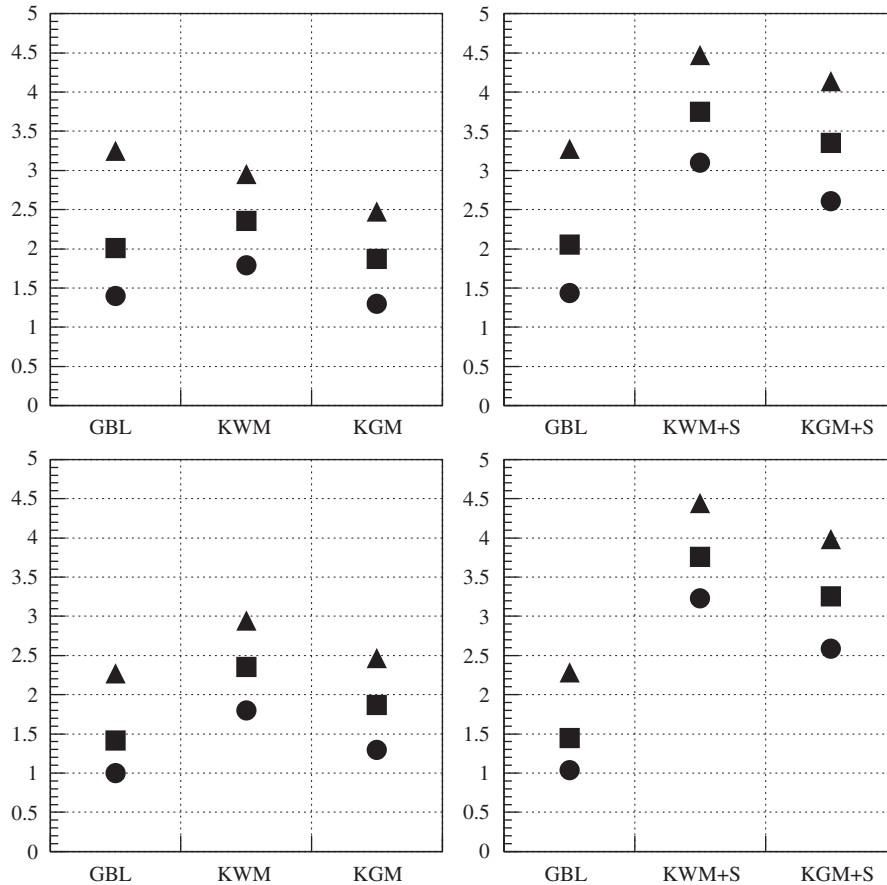


Fig. 1. Computing time for track refitting using different algorithms, propagation Jacobians and setups of scatterers normalized to the fastest case. The top plots are with the detailed helix propagator, the bottom ones with a simplified version, the left ones for filtering (one track end) only and the right ones for filtering and smoothing (results at both ends). The methods shown are general broken lines (GBL) and Kálmán filtering in weighted mean (KWM) or gain matrix (KGM) formalism. The case of the scattering planes coinciding with the measurement planes is indicated by circles, one thin scatterer between measurements by squares and two thin scatterers between measurements by triangles.

dominated by the inversion of 5×5 matrices ($\mathcal{O}(5^3)$ operations, \mathbf{C} in Eq. (5b)) for the addition of a single measurement. Using the gain matrix formalism requires the inversion of matrices of the size of the dimension d of the measurements ($\mathcal{O}(d^3)$ operations) and several multiplications of $5 \times d$ matrices ($\mathcal{O}(5^2 \cdot d), \mathcal{O}(5 \cdot d^2)$ operations). The smoothing step to get optimal local track parameters at all and not only the last point involves inversion and multiplications of 5×5 matrices ($\mathcal{O}(5^3)$ operations). For small d the gain matrix formalism should be fastest.

The general broken lines can add several (n , between one and all) measurements with one fit. The effort is defined by the construction of the linear equation system from the non-zero derivatives from the propagation to the previous and next thin scatterer and its solution using $\mathcal{O}(n_{\text{par}}) \cdot (m+1)^2$ operations with in general band width $m=5$ and $n_{\text{par}}=2n_{\text{scat}}+1$, depending on the number of thin scatterers $n_{\text{scat}}=2, \dots, 2n$.

In a toy setup the performance of track refitting has been compared. Tracks in a detector consisting of 10–50 planes with two independent measurements ($d=2$) and thin scatterers coinciding with the measurement planes (simple model for a silicon tracker) or one or two thin scatterers between adjacent planes have been simulated. The magnetic field \mathbf{B} is assumed to be constant between adjacent points and the initial trajectory uses a detailed helix propagator [8] or the simplified version for weak deflection (limit $|\mathbf{B}|/p \rightarrow 0$) as in Ref. [6]. Starting from the initial trajectory corrections and covariance matrix of track parameters at one or both track ends have been calculated. The results of this test are shown in Fig. 1. For fit results at one track end Kálmán filtering is compared with a general broken lines fit including one transformation from fit to local track parameters. Depending on the number of scatterers per measurement the fit with general broken lines is almost as fast as the Kálmán filter for the detailed propagator ($\pm 25\%$) and faster for the simplified one (10–80%). This speedup is due to the reduced number of non-zero elements in \mathbf{H} for the matrix construction and reduced band width $m=4$. For the Kálmán filter the gain matrix formalism is a little faster than the weighted mean formalism. For fit results at both track ends Kálmán filtering and smoothing are compared with a general broken lines' fit including two transformations from fit to local track parameters. In this configuration the general broken lines are faster in all cases tested (30–300%).

5. Summary

A trajectory based on general broken lines (GBL) is a track refit to add the description of multiple scattering to an initial trajectory based on the propagation in a magnetic field (and average energy loss). The initial trajectory can be the result of a fit of the

measurements (internal seed) or a prediction from another detector part (external seed). It is constructed from a sequence of (pairs of) thin scatterers describing the multiple scattering in the material between adjacent measurement planes. Predictions for the measurements are obtained by interpolation between the enclosing scatterers and triplets of adjacent scatterers define kink angles with variance according to the material of the central scatterer. The required propagation (on the initial trajectory) from a measurement plane or scatterer to the previous and next scatterer is using locally a linearized track model. This linearization may necessitate iterations of the fitting procedure. The matrix of the resulting linear equations system is a bordered band matrix allowing for a fast solution by a root-free Cholesky decomposition in a time proportional to the number of measurements. At all scatterers and measurement planes corrections for the local track parameters are determined. The corresponding covariance matrices and the pulls for the measurements and kinks require only the calculation of the bordered band part of the inverse matrix.

In a simulated detector, track refitting with GBL is about as fast or a little faster than Kálmán filtering, depending on the multiple scattering setup and on the complexity of the propagation Jacobian. In comparison with Kálmán filtering and smoothing it is up to a factor three faster.

A GBL fit with external seed and one (additional) measurement is equivalent to the filtering step of the track fit with a Kálmán filter.

As the track refit can provide the complete covariance matrix of all track parameters GBL are well suited as track model for alignment and calibration with Millepede-II [4].

References

- [1] A. Strandlie, R. Frühwirth, Reviews of Modern Physics 82 (2010) 1419.
- [2] P. Billoir, Nuclear Instruments and Methods in Physics Research Section A 225 (1984) 352.
- [3] R. Frühwirth, Nuclear Instruments and Methods in Physics Research Section A 262 (1987) 444.
- [4] V. Blobel, Nuclear Instruments and Methods in Physics Research Section A 566 (2006) 14.
- [5] V. Blobel, Nuclear Instruments and Methods in Physics Research Section A 566 (2006) 14.
- [6] V. Blobel, C. Kleinwort, F. Meier, Computer Physics Communications (2011). doi:10.1016/j.cpc.2011.03.017.
- [7] C. Amsler, et al., Particle Data Group, Physics Letters B667 (2008) 1 (2009 partial update for the 2010 edition).
- [8] A. Strandlie, W. Wittek, Nuclear Instruments and Methods in Physics Research Section A 566 (2006) 687.
- [9] I.S. Duff, A.M. Erisman, J.K. Reid, Direct Methods for Sparse Matrices, Oxford Science Publications, Oxford, 1986.
- [10] W.D. Hulsbergen, Nuclear Instruments and Methods in Physics Research Section A 600 (2009) 471.